NTP
National Toxicology Program

Review of the
Biomolecular Screening Branch
by the
NTP Board of Scientific Counselors

Poster Session

November 30, 2010

Rall Building

National Institute of Environmental Health Sciences

Research Triangle Park, NC

This page left intentionally blank

Table of Contents

Category	Authors	Title	Page #
Tox21 Partners-EPA	Judson et al.	EPA NCCT Databases and Knowledgebases: ACToR, DSSTox, ToxRefDB, ToxMiner, ExpoCastDB, and VT-KB	1
Tox21 Partners-EPA	Reif et al.	Chemical Prioritization and Profiling Applications of ToxPi for Tox21 Data	2
Tox21 Partners-EPA	Shah et al.	The EPA Virtual Liver Project	3
Tox21 Partners-FDA	Bearden et al.	Study of The FDA-Approved Drugs for Identification of *In Vitro* Biomarkers Related to Drug –Induced Liver Injury	4
Tox21 Partners-FDA	Benz et al.	U.S. FDA CDER's Toxicological and Adverse Human Clinical Effect QSAR Models	5
Assays & Pathways	Ramos et al.	Assessment of Mitochondrial Toxicity of Environmental Chemicals using a Quantitative High Throughput Screening Approach	6
Assays & Pathways	Wang et al.	Development and Validation of a Cell-Based Assay to Identify Small Molecule Modulators of Tristetraprolin, a Critical Regulator of Inflammation	7
Assays & Pathways	Xia et al.	Evaluation of Environmental Chemicals on the Blockage of hERG Channels	8
Informatics	Goldsmith et al.	To Hit or Not To Hit?: In Silico Models of In Vitro Nuclear Receptor Transactivation	9
Informatics	Huang et al.	Chemical Genomics Profiling of Environmental Chemical Modulation of Human Nuclear Receptors	10
Informatics	Lim et al.	Preliminary Test Estimation Procedures in High Throughput Screening Assays	11
Informatics	Myatt and Cross	An Integrated Prediction System To Support Toxicology Assessment	12
Informatics	Rashid et al.	NTPCEBS: A Repository for Tox21 qHTS Data	13
Informatics	Shockley et al.	A Decision Tree Algorithm for Analyzing Quantitative High Throughput Screening Data	14
Targeted Testing	Chan et al.	Evaluating Tools and Models used for Quantitative Extrapolation of *In Vitro* High Throughput Screening Data to Risk Management Decisions	15

Category	Authors	Title	Page #
Activities	Boyd et al.	Analysis of the ToxCast 320 Pesticide Library using a *Caenorhabditis elegans* growth Assay	16
Activities	Snyder et al.	Development of a *Caenorhabditis elegans* Movement Assay for Assessing Neurotoxicity	17
Activities	Merrick et al.	Mining the NTP Archives for Gene Signatures	18
Activities	Auerbach et al.	Acquisition of Drugmatrix® for the NTP and Tox21	19
Activities	French et al.	Benzene ADME and Toxicokinetics in Multiple Mouse Strains and Targeted Testing Strategies	20
Activities	French	Systems Analysis of Benzene Toxicity and Associated Response Pathways	21

Poster 1

EPA NCCT DATABASES AND KNOWLEDGEBASES: ACTOR, DSSTOX, TOXREFDB, TOXMINER, EXPOCASTDB, VT-KB

Richard Judson, Ann Richard, Matthew Martin, Elaine Cohen Hubal, Sumit Gangwal, Imran Shah, Thomas Knudsen, Marti Wolf, Doris Smith, James Vail, Shad Mosher, Tom Transue, Tommy Cathey, and Amar Singh

National Center for Computational Toxicology, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

The EPA National Center for Computational Toxicology is developing a set of databases and knowledgebases to support research in toxicology and to enhance dissemination of environmental information and data transparency. This poster summarizes our efforts in this area. ACToR (Aggregated Computational Toxicology Resource) is a set of databases compiling information on chemicals in the environment from a large number of public and in-house EPA sources. DSSTox is focusing on curation of chemical structures and associated toxicologically-relevant data sets. ToxRefDB is compiling detailed information from animal studies (sub-chronic, chronic, reproductive, developmental). ToxMiner is housing data from the ToxCast™ and Tox21 programs and related annotations. ExpoCastDB is compiling detailed exposure data from EPA and external sources. VT-KB (Virtual Tissue Knowledgebase) is compile information from the open literature linking chemicals, genes, pathways and disease. Currently, ACToR, ToxRefDB and DSSTox are publicly available, and the remaining resources are moving towards public release.

This abstract does not necessarily reflect U.S. EPA policy.

Poster 2

CHEMICAL PRIORITIZATION AND PROFILING APPLICATIONS OF TOXPI FOR TOX21 DATA

David Reif, Matt Martin, Shirlee Tan, Keith Houck, Richard Judson, Ann Richard, Thomas Knudsen, Ander Wilson, Sumit Gangwal, Elaine Cohen Hubal, David Dix, and Robert Kavlock

National Center for Computational Toxicology, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

The Toxicological Prioritization Index (ToxPi™) framework was developed as a decision-support tool to aid in the rational prioritization of chemicals for integrated toxicity testing. ToxPi consolidates information from multiple domains—including ToxCast™ *in vitro* bioactivity profiles (a wide-ranging battery of over 500 high throughput screening assays), inferred toxicity pathways, exposure predictions, and chemical properties/descriptors—into comprehensive toxicity scores and multivariate visualizations representing the contribution of each data domain to overall priority rankings. Here, we demonstrate applications of ToxPi that integrate data from across the Tox21 partnership. These include an application for endocrine activity, an application that simultaneously considers four sectors of toxicological concern (systemic, cancer, developmental, and reproductive), and an application that foresees how ToxPi could be used for assigning future test chemicals to toxicological activity "neighborhoods". Taken together, these applications demonstrate the utility of this framework for communicating results of the high-dimensional data generated as part of Tox21 and supporting decisions on targeted testing needs.

This abstract does not necessarily reflect U.S. EPA policy.

Poster 3

THE EPA VIRTUAL LIVER PROJECT

Imran Shah and John Wambaugh

National Center for Computational Toxicology, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

The objective of the Virtual Liver (v-Liver) is to provide information for tiered toxicity testing. The project is developing *in silico* decision support tools for analyzing the potential risk of toxic effects from environmental chemicals including identification of mode of action and characterization of human relevance. These tools leverage available *in vitro* High Throughput Screening (HTS) and High-Content Screening (HCS) data such as -*omics* data, and public domain sources of biological information to develop computational systems models for predicting toxicity and prioritizing chemicals for further testing. Systems models complement traditional and HTS-based approaches to hazard identification through data-driven mechanistic models of toxicity pathways to enable rapid development of quantitative predictive models that can be used to prioritize chemicals based on the potential for liver toxicity.

The v-Liver is part of a broader EPA effort on Virtual Tissues (VT) aimed at reducing the magnitude and spectrum of animal testing by integrative *in silico* and *in vitro* models, which recapitulate the properties of intact organs. The other VT project includes the Virtual Embryo (v-Embryo) for investigating developmental toxicity. The modular framework of these innovative computational tools can be broadly applied to mining information about signalling and metabolic pathways from literature and databases, a knowledgebase for organizing evidence about 'toxicity pathways', statistical tools to analyze HTS-HCS data, and systems modelling tools to estimate quantitative dose-response.

This poster will present v-Liver tools for evaluating a subset of environmental chemicals for liver cancer. First, we use public domain biological information to analyze potential toxicity pathways in liver cancer. Second, we include HTS assays from ToxCast™ to analyze how chemicals perturb these pathways based on potency and dosimetry and compare these findings with rodent toxicology data from ToxREFDB. Third, we describe efforts to simulate the dose- and time-dependent perturbations of 22 chemicals on these pathways using a novel systems modelling framework. The v-Liver framework aims to integrate these tools into an interactive collaborative workflow that synthesizes expert knowledge, public domain data, HTS and rodent toxicology to address risk assessment goals.

This abstract does not necessarily reflect U.S. EPA policy.

Poster 4

STUDY OF THE FDA-APPROVED DRUGS FOR IDENTIFICATION OF *IN VITRO* BIOMARKERS RELATED TO DRUG–INDUCED LIVER INJURY

Edward Bearden[1], Menghang Xia[2], Minjun Chen[1], Ruili Huang[2], Sampada Shahane[2], Min Zhang[1], Jie Zhang[1], Qiang Shi[1], Vikrant Vijay[1], Christopher P. Austin[2], and Weida Tong[1]

[1] National Center for Toxicological Research, U.S. Food and Drug Administration, Jefferson, AR, USA

[2] NIH Chemical Genomics Center, National Institutes of Health, Bethesda, MD, USA

A significant percentage of drugs that are withdrawn from the market or fail during the clinical trial stage of development do so because of liver toxicity. Drug induced liver injury (DILI) has been identified by the FDA as a key area for development of better evaluation tools and safety biomarkers. Conventionally, liver toxicity has been investigated using a set of standardized animal-based studies, which are very expensive and time-consuming, and often fail to predict liver toxicity in human. In addition, they often do not provide detailed mechanistic information that can allow animal-to-human extrapolation. An alternative approach which is being widely evaluated is to prioritize and/or supplement animal testing with a battery of mechanistically informative *in vitro* assays. In this project, we developed predictive models using *in vitro* data to identify mechanistically relevant biomarkers for the FDA-approved drugs. Specifically, we applied a battery of cell-based assays in a quantitative high throughput screening (qHTS) format developed by the NIH Chemical Genomics Center (NCGC) for the drugs studied in the FDA's Liver Toxicity Knowledge Base (LTKB) project. We then developed computational models based on the qHTS assay results for these drugs. The outcome of the project will provide a wealth of mechanistic information and biomarker models for liver toxicity. In this poster, the preliminary results of this project will be presented. We anticipate that the project will provide a resource for the FDA to utilize and reference when liver toxicity issues arise during various stages of the regulatory review process. The project will also serve as a proof-of-concept approach for developing knowledge and resources related to other drug safety issues (e.g., cardiac disease) essential to the FDA's regulatory mission.

Poster 5

U.S. FDA CDER'S TOXICOLOGICAL AND ADVERSE HUMAN CLINICAL EFFECT QSAR MODELS

R. Daniel Benz[1], Naomi L. Kruhlak[1], and Barbara L. Minnier[1,2]

[1]FDA Center for Drug Evaluation and Research (CDER), Silver Spring, MD, USA
[2]GlobalNet Services, Inc., Rockville, MD, USA

The U.S. FDA Center for Drug Evaluation and Research (CDER) Office of Testing and Research (OTR) computational toxicology group creates databases of toxicological and clinical studies for the development of (quantitative) structure-activity relationships (SAR). We employ research collaboration agreement (RCA) leveraging mechanisms to provide to all interested users (Q)SAR models based on these databases, and to assist (Q)SAR software providers with the development of a variety of computational software programs that offer predictive (Q)SAR modeling capabilities (MC4PC, Predictive Data Miner, Derek for Windows, BioEpisteme and SciQSAR). We are also involved in initiatives to harvest and publicly share toxicity data that support our modeling activities, and promote the use of computational techniques that allow scientific knowledge to be extracted from proprietary chemical structures without revealing their identities. Recently, we have expanded our computational toxicology capabilities by creating new comprehensive databases and (Q)SAR model suites describing adverse human health effects and the development of a weight-of-evidence approach for interpreting (Q)SAR predictions from multiple models and software platforms. In addition, our ongoing research program continues to upgrade our existing databases and models with re-interpretations of older data and new studies as they become available, and identifies new areas of interest, such as the (Q)SAR modeling of phospholipidosis. Our *in silico* tools are used within FDA/CDER to estimate the potential toxicity of chemicals of interest, such as drug metabolites and contaminants, and includes the toxicological profiling of chemicals for which little or no traditional test data are available or for which results are equivocal. The work described here supports FDA's Critical Path Initiative which encourages the use of computational tools in a regulatory environment to enhance the approval process for FDA-regulated products.

Poster 6

ASSESSMENT OF MITOCHONDRIAL TOXICITY OF ENVIRONMENTAL CHEMICALS USING A QUANTITATIVE HIGH THROUGHPUT SCREENING APPROACH

Matias S. Attene Ramos[1], Ruili Huang[1], Srilatha Sakamuru[1], Sampada Shahane [1], Louie Shou[1], Kristine L. Witt[2], Raymond R. Tice[2], Christopher P. Austin[1], and Menghang Xia[1]

[1]NIH Chemical Genomics Center (NCGC), National Institutes of Health, Bethesda, MD, USA
[2]National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

As part of the U.S. Tox21 initiative, the NCGC is developing and optimizing cell-based and biochemical assays suitable for quantitative high throughput screening (qHTS) in 1536-well format. This effort will generate pathway profiles for environmental compounds that will facilitate the evaluation of mechanisms of toxicity and prioritization for more extensive testing, as well as the development of predictive models for *in vivo* toxicity. In this study, we optimized a mitochondrial membrane potential assay using the water-soluble JC-10 sensor to detect mitochondrial depolarization in HepG2 cells and we then used this method to evaluate the mitochondrial toxicity of 1408 environmental compounds provided by the NTP. In response to mitochondrial depolarization, the ratio of the cytosolic green fluorescent monomeric form to the mitochondrial red fluorescent aggregate form increases. Of the 1408 compounds screened over a 14-point concentration curve (0.5 nM to 92 μM), 44 compounds disrupted the mitochondrial potential in HepG2 cells after treatment for one hour. We selected 33 compounds for further studies, including high-content imaging. Thirty-two compounds were confirmed in both fluorescence plate reader and imaging assay formats. To study the structure-activity relationship of these mitochondrial disruptors, we clustered these compounds by structural similarity. This analysis resulted in four structural clusters and 15 singletons. These clusters may be useful for identifying structural features associated with mitochondrial toxicity. Our results confirm the robustness of this assay for identifying mitochondrial membrane-potential disruptors in qHTS format. Supported by NIEHS Interagency Agreement Y2-ES-7020-01.

Poster 7

DEVELOPMENT AND VALIDATION OF A CELL-BASED ASSAY TO IDENTIFY SMALL MOLECULE MODULATORS OF TRISTETRAPROLIN, A CRITICAL REGULATOR OF INFLAMMATION

Hang Wang[1], Ruili Huang[1], James Inglese[1], Christopher P. Austin[1], Thierry J. Horner[2], Betsy A. Kennington[2], Keith R. Shockley[3], Raymond R. Tice[3], Ronald L. Johnson[1], and Perry J. Blackshear[2]

[1] NIH Chemical Genomics Center, National Institutes of Health, Bethesda, MD, USA

[2] Laboratory of Signal Transduction, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[3] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

A number of disease pathologies are mediated by excessive levels of tumor necrosis factor alpha (TNF), including septic shock, rheumatoid arthritis, Crohn's disease, psoriasis and inflammation linked with cancer. Tristetraprolin (TTP) is a key endogenous protein regulator of inflammation that acts by destabilizing select cytokine mRNAs such as that encoding TNF. The goal of this project is to identify small molecules that modulate TTP protein accumulation in macrophages, the major cellular producer of TNF during inflammation. Compounds that stimulate TTP biosynthesis in macrophages should lead to destabilization of the TNF mRNA and thereby have an anti-inflammatory effect. Alternatively, environmental compounds that inhibit TTP biosynthesis in macrophages should result in greater stability of the TNF mRNA and thereby exacerbate inflammation. To identify compounds that modulate TTP levels, we developed a RAW 264.7 mouse macrophage cell line that stably expresses a human TTP-GFP (green fluorescent protein) reporter fusion protein (TTP cells). The parental RAW 264.7 cells exhibit increased TTP mRNA and protein expression in response to LPS stimulation, the same response as seen in primary mouse macrophages. This reporter contains the human TTP promoter, both exons and the single intron, with eGFP fused in-frame to the final TTP exon to express a full-length TTP-eGFP fusion protein. Regulated expression of the fusion protein was documented by immunofluorescence and western blotting. For high throughput screening, TTP-GFP expression was detected by a laser scanning microplate cytometer to enumerate GFP positive cells. The assay was miniaturized into the 1536-well plate format and validated in a titration-response screen of the LOPAC1280 collection at eight concentrations, using LPS (lipopolysaccharide) as a positive control. We also screened the NTP 1408 compound library in this assay in agonist and antagonist modes. From this screen, active compounds were identified based on automated curve fitting and classification of the concentration-response data. These results demonstrate that the TTP induction assay can identify stimulators of human TTP expression in macrophages, potentially leading to the development of anti-inflammatory drugs. In addition, it can identify compounds in the environment that inhibit this important innate immune defense mechanism. Supported by the NIH Intramural Research Program and NIEHS Interagency Agreement Y2-ES-7020-01.

Poster 8

EVALUATION OF ENVIRONMENTAL CHEMICALS ON THE BLOCKAGE OF hERG CHANNELS

Menghang Xia[1], Sampada Shahane[1], Ruili Huang[1], Steve Titus[1], Enoch Shum[2], Yong Zhao[2], Noel Southall[1], Wei Zheng[1], Kristine L. Witt[3], Raymond R. Tice[3], and Christopher P. Austin[1]

[1] NIH Chemical Genomics Center, National Institutes of Health, Bethesda, MD, USA
[2] Cerep, Inc., Redmond, WA, USA
[3] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

The human ether-a-go-go-related gene (hERG) channel, a member of a family of voltage-gated potassium (K^+) channels, plays a critical role in the repolarization of the cardiac action potential. The reduction of hERG channel activity as a result of adverse drug effects or genetic mutations may cause QT interval prolongation and potentially lead to acquired long QT syndrome. Thus, screening for hERG channel activity is important in drug development. Cardiotoxicity associated with the inhibition of hERG channels by environmental chemicals is also a public health concern. To assess the inhibitory effects of environmental chemicals on hERG channel function, we screened the National Toxicology Program (NTP) collection of 1408 compounds by measuring thallium influx into cells through hERG channels. A group of compounds with hERG channel inhibition were identified with IC_{50} potencies ranging from 0.26 to 22 μM. Most of these compounds were confirmed as hERG channel blockers in an automated whole cell patch clamp experiment. In addition, we discovered that a quaternary ammonium compound (QAC) series containing seven compounds had the structure-activity relationship on hERG channel inhibition. Among these compounds, tetra-n-octylammonium bromide was the most potent with an IC_{50} value of 260 nM in the thallium influx assay and 80 nM in the patch clamp assay. Our study demonstrated that profiling environmental compound libraries for hERG channel inhibition provides information useful in prioritizing these compounds for cardiotoxicity assessment *in vivo*.
Supported by NIEHS Interagency Agreement Y2-ES-7020-01.

Poster 9

"TO HIT, OR NOT TO HIT?": IN SILICO MODELS OF *IN VITRO* NUCLEAR RECEPTOR TRANSACTIVATION

Michael-Rock Goldsmith[1], Daniel T. Chang[1], and Raymond R. Tice[2]

[1] National Exposure Research Laboratory, US Environmental Protection Agency, Research Triangle Park, NC, USA
[2] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

In silico models are used to inform and prioritize bioassay requirements. Here, we develop a model for predicting the ability of compounds to activate estrogen/androgen/ thyroid (EAT) receptors in HEK293 cell line transactivation assays, based on data obtained on 309 unique chemicals in ToxCast™ Phase 1. In our model, a hit for either "E" or "A" or "T" or any combination/permutation, in both or either agonist/antagonist mode according to the assays performed by the NIH Chemical Genomics Center (NCGC) is a "hit". A "non-hit" is a chemical that shows no activity on any of the same targets with the same ToxCast™ training set. A decision tree classifier was developed with our functional definition of *in vitro* transactivation "hit" described above as our "class field", and a total of 205 assay endpoints (43 Novascreen *in vitro* assays, 150 molecular docking assays, and 12 QSPR [Quantitative Structure-Property Relationship]/ADME QSAR [absorption, distribution, metabolism, and excretion; quantitative structure-activity relationship) were selected as "tree fields". Our classification tree was developed using two-fold cross validation using the binary decision tree classifier implemented in Molecular Operating Environment (Chemical Computing Group, Montreal QC Canada). The most valuable "bits" that provided information that could discriminate, hence enrich hits to non-hits were MDCK (Madin-Darby canine kidney cell) permeability, 1GS4 (androgen receptor) and 2UW9 (pkb or akt1) binding (*in silico* molecular docking). An interesting point is that the QSAR-predicted MDCK permeability was the most useful parameter in enriching a dataset. Another interesting feature is how the *in vitro* and additional QSPR and docking results were neglected, and only the information bits with maximum ability to enrich hit/no-hit dataset were preserved; a cell permeability property and two discrete docking targets; the resulting accuracy of the model has a misclassification rate of ~11% (based on the chemical space of the training chemicals). These methods assist in the interpretation of the actual assay and outline key determinants of transactivation *in vitro*: cell permeability, nuclear receptor binding, and specificity of competing targets; a nuclear translocatable "AR" protein over obligate cytoplasm confined non–specific protein. Applied to the NCGC Phase I Tox21 2900 compound library, a total of 676 chemicals were "predictable" based on the domain of applicability, of which ~ 65 were considered to be "active" based on modeled receptor specificity and cell permeability. In summary, our *in silico* model provides us with a mechanism of identifying the most likely chemical candidates that have the proper trans-cellular absorption criteria in addition to identifying fundamental ligand/biomolecular interaction determinants encoded *in vitro* or *in silico* (docking) results that give rise to segregated and enriched data. *This abstract does not necessarily reflect U.S. EPA policy.*

Poster 10

CHEMICAL GENOMICS PROFILING OF ENVIRONMENTAL CHEMICAL MODULATION OF HUMAN NUCLEAR RECEPTORS

Ruili Huang[1*], Menghang Xia[1], Ming-Hsuang Cho[1], Srilatha Sakamuru[1], Paul Shinn[1], Keith A. Houck[2], David J. Dix[2], Richard S. Judson[2], Kristine L. Witt[3], Robert J. Kavlock[2], Raymond R. Tice[3], and Christopher P. Austin[1]

[1] NIH Chemical Genomics Center, National Institutes of Health, Bethesda, MD, USA

[2] National Center for Computational Toxicology, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

[3] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

The large and increasing number of chemicals released into the environment demand more efficient and cost effective approaches for assessing environmental chemical toxicity. The U.S. Tox21 program has responded to this challenge by proposing alternative strategies for toxicity testing, among which the quantitative high throughput screening (qHTS) paradigm has been adopted as the primary tool for data generation on large chemical libraries against a wide spectrum of assays. This will serve as the basis for predictive model building, hypothesis generation, and prioritization for further testing. To realize the goals set forth by Tox21, selection of appropriate assays for screening and the quality of data generated from these assays are of central importance. As a proof-of-concept study, the Tox21 pilot phase collection of approximately 2,900 compounds was profiled against a panel of ten human nuclear receptors in qHTS format. We have evaluated these data from a chemical genomics point of view, in addition to assessing data reproducibility as a measure of data quality and applicability for down stream analyses. This exercise has enabled us to formulate data driven strategies for discriminating true signals from artifacts, assay selection, and prioritization. Analyses of connections from protein sequences of the receptors to compound phenotypes and structures have shown the nuclear receptor profiling data to be a valuable data set for these purposes. However, interference from assay artifacts, most prominently cytotoxicity in antagonist mode assays, remains a challenge calling for better solutions based on more in-depth analysis of existing data, and additional data generated on more compounds and assays. Supported by NIEHS Interagency Agreement Y2-ES-7020-01.

Poster 11

Preliminary Test Estimation Procedures in High Throughput Screening Assays

Changwon Lim[1], Pranab K. Sen[2,3], and Shyamal D. Peddada[1]

[1] Biostatistics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[2] Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, 338 Hanes Hall, CB#3260, Chapel Hill, NC, USA

[3] Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Quantitative high throughput screening (qHTS) assays use cells or proteins to screen thousands of compounds in a short period of time. Data generated from qHTS assays are then fitted using a nonlinear regression model and decisions regarding the activity of a chemical are made using the estimates of the parameters of the model. For such data sets, the error variance may be constant across dose groups (homoscedasticity) or vary depending upon the dose (heteroscedasticity). Because thousands of compounds are evaluated in qHTS assays based on the estimates of the parameters of nonlinear models, it is important to apply an estimation procedure which is robust to the error variance structure. Additionally, it is desirable for the procedure to be robust to outliers and influential observations. In this poster we describe a preliminary test estimation (PTE)-based methodology for drawing inferences regarding the parameters of a Hill model with application to qHTS assays. Performance of the PTE based methodology, in terms of false discovery rate (FDR) and power, is evaluated using a simulation study mimicking real qHTS data. Results suggest that the proposed methodology controls the FDR while having larger power than some of the currently used methodologies. Using a data set obtained from the National Toxicology Program, we illustrate the proposed methodology.

Poster 12

An Integrated Prediction System to Support Toxicology Assessment

Glenn J. Myatt and Kevin P. Cross

Leadscope, Inc., 1393 Dublin Rd, Columbus, OH, USA

Assessing the safety of chemicals is critical for NIEHS in its missions to reduce the burden of environmentally associated disease and dysfunction, and promote the public's right to a healthy, quality environment. In support of this mission, new technologies are being developed to rapidly assess the safety of chemicals. In collaboration with other government agencies through Tox21, many sources of disparate information are coming online including human adverse events, studies using laboratory animals, qHTS (quantitative high throughput screening), gene expression, and biological pathways data. A new multi-tier client/server/database system is being developed through a phase I NIEHS Small Business Innovative Research (SBIR) grant that will support integrating these diverse sources of data. These data will be augmented with online toxicity predictive model results from several sources including Leadscope's and MultiCASE's QSAR (quantitative structure-activity relationship*)* models and Derek structural alerts from Lhasa Limited. In addition, a number of simple tools are proposed to help scientists make sense of the information. Methods to help prioritize chemicals for more extensive testing based on similar profiles, such as from qHTS or gene expression data, to a reference chemical are being explored. In addition, tools to support understanding mechanisms of action, human relevant predictions, and NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) validation studies are also being investigated. An important aspect of the proposed software is that it should be available through web browsers on different operating systems and it should support scientists with diverse experience and training as well as integrating with current NIEHS' workflows and systems. Supported by SBIR Contract #HHSN273201000006C.

Poster 13

NTPCEBS: A REPOSITORY FOR TOX21 QHTS DATA

Asif Rashid[1], Hui Gong[1], and Jennifer Fostel[2]

[1] NTP Computer and User Support, Team Vistronix, Research Triangle Park, NC, USA
[2] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

NTPCEBS (National Toxicology Program Chemical Effects in Biological Systems) is a database housing data of interest to toxicologists. NTPCEBS will be the repository for the legacy data from the NTP (National Toxicology Program) as well as data from ongoing studies by the NTP, including the Tox21 effort. At present NTPCEBS houses Tox21 data from the qHTS (quantitative high throughput screening) screening effort carried out by the NTP and NCGC (National Chemical Genomics Center). Assay data are obtained from the NCGC, checked for format and data integrity, then metadata describing the assay are added to the file and it is loaded into NTPCEBS. In addition to the normalized data, NTPCEBS houses multiple analysis paths for each assay, using methods derived at the NCGC and NTP to identify active and inactive compounds in particular assays. The user can query for compounds that are active (or inactive) in a particular assay. Alternatively the user can identify assays in which a particular compound is active (or inactive). These queries each generate a “hits list”. NTPCEBS permits the user to view the original data and the curve fit for each compound/assay in the hits list. The user can run multiple queries and then combine or intersect hits lists. The user can then download lists of chemicals or assays with the desired activity. NTPCEBS provides a view of the data in which the results from all assays and compounds are ordered using a particular analysis method to assess activity, so that the user can see the totality of activity across all the qHTS data in NTPCEBS before retrieving data from a single assay or compound. The user can also use this view to identify assays or chemicals of interest. Data from the first phase of screening are available in CEBS (http://cebs.niehs.nih.gov/) under the “HTS” tab. Once the next phase of Tox21 is underway these data will also be housed in NTPCEBS and released once the data have been validated for quality.

Poster 14

A DECISION TREE ALGORITHM FOR ANALYZING QUANTITATIVE HIGH THROUGHPUT SCREENING DATA

Keith R. Shockley[1], Grace E. Kissling[1], Ruili Huang[2], Menghang Xia[2], Christopher P. Austin[2], and Raymond R. Tice[1].

[1] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[2] NIH Chemical Genomics Center, National Institutes of Health, Bethesda, MD, USA

The ability of a substance to induce a toxicological response is better understood by analyzing the response profile over a broad range of concentrations (or doses) rather than by evaluating effects that occur at a single concentration (or dose). *In vitro* quantitative high throughput screening (qHTS) assays are multiple-concentration experiments that play an important role in NTP's efforts to advance toxicology from a predominantly observational science at the level of disease-specific models to a predominantly predictive science based on broad inclusion of target-specific, mechanism-based, biological observations. The analysis of qHTS data has largely been motivated by the conservative focus of pharmaceutical applications (i.e., minimizing the risk of Type I error) and generally has relied on heuristics rather than statistical tests to make activity calls. To evaluate the activity within qHTS studies, we developed a four-stage decision tree statistical model and applied it to normalized concentration-response data from eleven cell-based agonist nuclear receptor assays (androgen receptor, estrogen receptor, farnesoid X receptor, glucocorticoid receptor, peroxisome proliferator-activated receptor delta and gamma, retinoid X receptor, thyroid hormone beta receptor, vitamin D receptor, pregnane x receptor human and rat). Data were fit to a four-parameter Hill equation and an overall F-test comparing the best fit to the Hill equation and a horizontal line (no response) was calculated for each chemical. Substances with a robust dose-response were identified in the first stage. In the second stage, compounds not detected as active in the first stage were evaluated for a maximal response at the lowest dose by comparing the distribution of measured responses to a control value. Chemicals with a weak dose-response were identified in the third stage, and the final stage separated substances exhibiting a cytotoxic response at the lowest dose from inactive compounds. Our model identified more active compounds than a previously utilized heuristic approach.

Poster 15

EVALUATING TOOLS AND MODELS USED FOR QUANTITATIVE EXTRAPOLATION OF *IN VITRO* HIGH THROUGHPUT SCREENING DATA TO RISK MANAGEMENT DECISIONS

Melissa Chan, Xiaoqing Chang, and Michael DeVito

National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

There are a number of risk management decisions, which range from prioritization for testing to quantitative risk assessments. The utility of high throughput screening (HTS) data in these decisions depends on our understanding of how well these *in vitro* assays qualitatively and quantitatively extrapolate to *in vivo* responses. Extrapolation of *in vitro* data to *in vivo* responses comprises of both pharmacodynamic and pharmacokinetic extrapolations. Initial efforts comparing HTS data to human exposures have assumed that media concentrations can be considered equivalent to steady-state blood concentrations. For this to be true, then the partitioning of the chemical between the media and the cell should be equivalent to the partitioning of the chemical between the blood and the tissue. Using data from the literature, the relationship between *in vitro* effects and *in vivo* responses was evaluated using deltamethrin and permethrin as test chemicals. *In vitro* data on chemical-induced decreases in cell firing in rat primary hippocampal cell cultures were compared to *in vivo* effects on decreased motor activity. Using media concentrations as equivalent to blood concentrations indicates that the media concentrations at the NOEL (no observable effect level), ED50 (medium effective dose), and maximum response correspond well to blood concentrations of these pyrethroids at the equivalent effect level *in vivo*. However, when the data is compared on a cell or tissue concentration, the *in vitro* system requires 50 times more chemical in the cells compared to the brain in order to produce responses. These findings suggest that media concentration is not directly comparable to blood concentrations. This initial finding focused on two neuroactive chemicals with the brain as the target tissue. It is uncertain whether these findings are applicable to other chemicals, *in vitro* systems and/or *in vivo* responses. We are presently evaluating the qualitative and quantitative relationship between *in vitro* primary hepatocyte models and *in vivo* responses using Ah receptor ligands and xenobiotic metabolizing enzymes as the test case. These efforts will provide an initial evaluation of the uncertainty in using HTS and *in vitro* data for risk management decisions ranging from prioritization for testing to quantitative risk assessments.

Poster 16

ANALYSIS OF THE TOXCAST 320 PESTICIDE LIBRARY USING A *CAENORHABDITIS ELEGANS* GROWTH ASSAY

Windy A. Boyd[1], Marjolein V. Smith[2], Julie R. Rice[1], and Jonathan H. Freedman[1, 3]

[1] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[2] SRA International, Durham, NC, USA

[3] National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

The National Toxicology Program, Environmental Protection Agency, and NIH Chemical Genomics Center are exploring the use of high-throughput *in vitro* tests and alternative model organisms to screen the toxicity of large numbers of chemicals to prioritize toxicity testing in traditional toxicological models. A number of endpoints of toxicity including reproduction, growth, and feeding have been developed for the nematode *Caenorhabditis elegans*. Using a COPAS Biosort, a flow cytometer designed to measure the size characteristics of individual nematodes, our lab group has partially automated these toxicity endpoints. In the *C. elegans* growth assay, the change in nematode size is measured after exposure to chemicals for 48h from the first to last larval stage (L1 - L4). The EPA's ToxCast 320 chemical library, which consists of mainly pesticide actives, was screened for effects on *C. elegans* growth first at a single high concentration (200 µM) and then at six lower concentrations (0.5, 1, 5, 10, 50, and 100 µM). A variety of statistical descriptors were used to characterize nematode responses to the chemicals. Descriptors include percent of control median size at each concentration, a distribution-free trend test, and comparison of estimated mean growth rates. At 200 µM approximately half of the chemicals were active causing severe decreases in *C. elegans* growth over 48h (i.e., the animals did not grow). A similar proportion of chemicals also caused a significant concentration-dependent decrease in growth from 0.5 to 100 µM. The ToxCast 320 library is composed of well-characterized chemicals with a large amount of toxicity data and has now been screened in hundreds of *in vitro*, biochemical, and alternative animal model assays. The summary statistics presented on the *C. elegans* growth data is additionally useful for comparison to results from rodent assays and from high throughput screens, allowing us to assess the usefulness of this assay in a tiered toxicity testing approach.

Poster 17

DEVELOPMENT OF A *CAENORHABDITIS ELEGANS* MOVEMENT ASSAY FOR ASSESSING NEUROTOXICITY

Daniel W. Snyder[1], Windy A. Boyd[1], Grace E. Kissling[1], and Jonathan H. Freedman[2]

[1] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[2]National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

Government agencies are using alternative toxicological model organisms to screen and prioritize chemicals to be used in traditional mammalian studies. The nematode *Caenorhabditis elegans* is a useful alternative model to characterize the behavioral effects of chemicals, because they possess a simple, well-characterized neuromuscular system. A movement assay has been developed to quantify *C. elegans* locomotion using a semi-automated motion tracking system. This system consists of an inverted fluorescence microscope equipped with a CCD camera, incubated motorized stage, and image analysis software. Using this system, a number of locomotion parameters were defined including curvilinear velocity, linearity, and amplitude of sinusoidal movement. RNA interference of a number of genes known to control normal nematode movement was used to evaluate the *C. elegans* movement assay. Changes in specific locomotion parameters indicated that the assay was capable of characterizing various loss-of-function phenotypes, which included uncoordinated and slow movement. A transgenic *C. elegans* strain expressing a mutant human tau protein, which is associated with neurodegenerative diseases such as Alzheimer's, was also analyzed. With the transgene expressed in all neurons, the mutant *C. elegans* moves slower than wild-type nematodes. Expression of the mutant tau transgene resulted in significantly impaired *C. elegans* movement, as determined by various motion parameters. The phenotype was rescued by six different genetic suppressors. Finally, the effects of three chemicals, cadmium, chlorpyrifos and tebuconazole, on *C. elegans* movement were quantified. All movement parameters significantly decreased by 1 mM cadmium and 0.1 to 10 μM chlorpyrifos. Tebuconazole, however, did not affect nematode locomotion at the highest concentration tested (1 mM). Together, these results confirm the utility of the *C. elegans* movement assay as a robust endpoint for investing the neurotoxicity of chemicals for toxicant screening and characterization.

Poster 18

Mining the NTP Tissue Archives for Gene Signatures

B. Alex Merrick

National Toxicology Program (NTP), National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

This initiative will test if gene expression signatures can be reliably derived from the NTP tissue archives. Signature expression profiles are critical sets of altered transcripts or proteins that distinguish toxicity and disease from a comparable normal state. Development of gene signatures can further our understanding of pathological changes, mechanisms, and critical pathways in chemical-induced toxicity. Further, such signatures could contribute to: the identification of useful targets for *in vitro* assays, an evaluation of the correlation between *in vitro* test results and *in vivo* toxicological outcomes, and the development of better predictive models of toxicity. Hundreds of thousands of tissues from the laboratory animals used in NTP's toxicological studies have been stored as formalin fixed, paraffin embedded (FFPE) blocks in the NTP archives. NTP toxicology studies sometimes also involve storing frozen tissues, but not always, so the NTP archives are comparatively limited in frozen tissue resources for extraction of full-length RNA for molecular profiling. Assuming that the suboptimal RNA that is extractable from FFPE tissues is an obstacle that can be overcome, molecular analyses on archival paraffin block tissues would greatly expand NTP's ability to link gene expression changes with disease outcomes while leveraging the considerable expense already invested in NTP toxicological studies. Several pilot studies have been initiated to evaluate different technologies for FFPE tissues to generate or validate gene signatures of agent-induced toxicity. The first such study determined that selective transcript amplification by quantitative polymerase chain reaction (qPCR) was indeed possible using RNA extracted from NTP archival FFPE rat liver tissues after subchronic aflatoxin B1 (AFB1) exposure. Gene fold changes observed were quantitatively comparable to those from Agilent microarrays. Second, ongoing studies using the same AFB1 liver and control tissues are evaluating global expression technologies to generate more informative signatures by RNASeq, a NextGen sequencing technology. A follow-up study on the same samples is underway in collaboration with Johns Hopkins University using a variation of RNASeq called 3'Seq. 3'Seq is expected to overcome limitations of degraded RNA in FFPE blocks, using livers from control and AFB1 treated animals. Third, the NTP Waalkes' laboratory and the BSB are developing NextGen sequencing methods to relate epigenetic changes in methylated DNA with transcript expression changes. Fourth, the quantitative nuclease protection assay or qNPA is a method for validation of gene signatures being piloted with NTP laboratory investigators. Fifth, a targeted proteomics platform is also being tested on fresh and FFPE samples for its ability to validate protein signatures. These pilot studies will guide NTP in use of promising technologies to derive or validate gene signatures from the NTP tissue archives that might be adapted into future HTS assays.

Poster 19

Acquisition of DrugMatrix® for the NTP and Tox21

Scott S. Auerbach

National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

DrugMatrix® is a toxicogenomics reference database and informatics system. The NTP has acquired DrugMatrix® as a unique and powerful resource to assist the scientific community in interpreting high throughput screening data for *in vivo* toxicological context and for better defining mechanisms of action of environmental toxicants. The three fold components of DrugMatrix® are a Toxicogenomics Database, informatics tools to query the database, and an extensive Tissue Library. Each component has several distinguishing features. The toxicogenomics database includes a graphics user interface that allows for rapid scoring of genomic signatures of toxicity and a framework that allows for the storage and analysis of multiple data types. The *in vivo* chemical exposure study data was based on treating male Sprague Dawley rats at multiple dose levels by oral gavage for acute, subacute and subchronic durations. These studies include extensive pharmacology, clinical chemistry, hematology, histology, body and organ weights, and clinical observations. *In vitro* data in rat primary hepatocytes involve assays of all toxicologically relevant endpoints and toxicogenomic (microarray) studies. Test agents include U.S. FDA approved drugs, standard biochemicals, and environmental toxicants with curation of all relevant public information. Advanced query tools allow access to all information in integrated data domains. Queries can combine literature-derived information on genes or compounds, pharmacological activity, expression changes, or chemical structure to identify agents that modulate, induce or repress specific genes. Affected biochemical pathways, specific mechanisms of action, or therapeutic class can be queried. Gene expression array experiments and the results of pharmacology panels can be queried in the DrugMatrix® database to identify compounds that elicit similar responses or structurally related reference compounds. The third component of DrugMatrix® is an extensive frozen Rodent Tissue archive that includes snap frozen tissues and corresponding total RNA from liver, heart, kidney, skeletal muscle, whole blood (or plasma) from rats treated with test agents. A series of tests have been designed to assess the quality and capabilities of DrugMatrix upon installation, including but not limited to an evaluation of Affymetrix®.CEL files for accepted quality control metrics, extraction of RNA from randomly selected frozen tissue samples for an assessment of integrity, examination of the database for demonstrated public access security controls, data upload architecture, and database security structure that ensures database integrity. The installation and assessment are currently in progress.

Poster 20

BENZENE ADME IN MULTIPLE MOUSE STRAIN RESULTS AND DEVELOPMENT OF TARGETED TESTING STRATEGIES

John E. French[1], Michael J. Cunningham[1], Gabriel A. Knudsen[2], Robert K. Kuester[2], I. Glenn Sipes[2]

[1] National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[2] University of Arizona, Tucson, AZ, USA

From a U.S. perspective, benzene is the most significant air toxic for which a cancer risk can be estimated. Based on a U.S. Environmental Protection Agency (EPA) National-Scale Air Toxics Assessment (see *http://www.epa.gov/ttn/atw/nata2002/*), exposure to benzene contributes ~25% of the average individual cancer risk. New risk assessment procedures using systems biology approaches were proposed in a recent EPA-sponsored meeting on *Advancing the Next Generation (NexGen) of Risk Assessment: The Prototypes*, which included new exposure assessment studies to determine inter-individual differences within susceptible populations exposed to benzene. Recent ADME (absorption, distribution, metabolism, excretion) studies in 18 strains of inbred mice show significant intra- and inter-strain variation in response to a single oral exposure to benzene (100 µg ^{14}C-benzene/kg body weight). Fifteen of these strains were re-sequenced in the NTP-Perlegen studies, where over 8 million SNPs were identified and the genomic structure of inbred laboratory mice revealed in great detail. Each of these individual inbred strains may be considered as an individual within a small population of genetically diverse mice. In this context, these studies are somewhat similar to identical twins in human populations, but the mice are homozygous at all loci whereas the human twins are exact copies of heterozygous individuals. Up to 10X differences for area under the curve (AUC) and clearance (CL_F) in blood levels and up to near 40X differences were observed for AUC and CL_F in the bone marrow in the 18 strains tested. C_{max} and T_{max} variation was not as great among these 18 strains. The significant difference observed in ADME between C57BL/6J and C3H/HeJ, the parental lines of the B6C3F1 inbred mouse is of significant interest. These data suggest the dramatic differences that may be observed due to gene-gene (protein-protein) interaction that affect quantitative measures of toxicity and the perturbation to pathways that regulate systems homeostasis. These data suggest that these individual differences indicate possible differences toxicokinetics that can be used in targeted testing strategies. We are using this approach to investigate the effects of a 28-day exposure to inhaled benzene at 0, 1, 10, or 100 ppm on hematotoxicity and genotoxicity in the strains used in the ADME studies. A twenty-eight day exposure was selected to achieve steady-state levels of pancytopenia and micronuclei induction. In this way, we attempt to reduce the within strain variation in response to benzene in order to determine the between strain variation in response. The results obtained will improve our ability to identify quantitative trait genes using haplotype association mapping and their human orthologs, which will enhance across species extrapolation for benzene cancer risk assessment.

Poster 21

SYSTEMS ANALYSIS OF BENZENE TOXICITY AND ASSOCIATED RESPONSE PATHWAYS

John E. French

National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC

Benzene is a ubiquitous environmental hemato- and genotoxicant. Benzene causes acute myeloid leukemia (AML) and myelodysplastic syndromes and has been associated with lymphoproliferative disorders including lymphoblastic disease in humans and mouse models of human cancer. Biological plausibility for a causal role of benzene in these diseases comes from its genotoxic effects and toxicity to hematopoietic stem cells or progenitor cells, from which leukemias and lymphomas arise. The effect of this toxicity in humans and rodent models is observed as a pancytopenia (i.e., reduction in red and white cell number) in the blood and blood forming organs (hematotoxicity), even in individuals exposed to very low levels of benzene in an occupational setting. Recent studies of low occupational benzene exposures have identified novel toxicogenetic and epigenetic biomarkers of exposure. To aid our understanding of the gene-environment interactions that may lead to benzene-induced cancer, it is necessary to understand benzene toxicity and hematopoietic disease as a quantitative trait, with interactions among many genes and at the systems level. Using the Comparative Toxicogenomic Database (CTD)[1] and the CTD knowledgebase (21 October 2010), benzene-chemical interactions were compiled from all hand curated publications showing a positive or negative correlation with a gene or genes that resulted in a reported biochemical reaction, chemical induced response, metabolic process, transcript expression, etc., without censorship of any data. The genes identified were analyzed using GeneGo's Metacore™[2] to identify 30 predicted networks and pathways. Significantly, the networks identified based upon the interacting genes and their associated ontology include, but were not limited to: (1) xenobiotic metabolism and cellular response to chemical exposure ($P = 3.5 \times 10^{-83}$), (2) cell cycle and regulation of cell and mitotic cycle ($P = 6.4 \times 10^{-45}$), (3) DNA damage, DNA recombination and DNA strand break repair ($P = 1.6 \times 10^{-56}$), (4) regulation of apoptosis ($P = 5.5 \times 10^{-33}$), and (5) cytokine mediated signaling and inflammation pathways ($P = 3.9 \times 10^{-30}$). Current efforts are focused on integrating these data (benzene-gene interactions and homeostasis pathways) with disease pathways (Genetic Association Database[3]) and comparing allele diversity in humans to those identified in our targeted testing strategy using inbred laboratory and wild-derived mice for haplotype association mapping of genes identified in low dose exposures to inhaled benzene. Using both databases and computational tools, we are working to identify those genes most critical to the benzene-induced disease phenotype and then to identify targets for validation and inclusion into high throughput systems assays for screening toxicants of human relevance.

---

[1] Comparative Toxicogenomics Database 2010; Nucleic Acids Research 37:D786-92, 2009; http://ctd.mdibl.org

[2] http://www.genego.com/metacore.php

[3] Genetic Association Database, *Nature Genetics* 2004, 36:431-432.